ISSN: 0974 – 3987
IJBST (2008), 1(3):10-16

Enhancing the Quality of Radiological Breast Cancer Images with Computer aided image segmentation technique

Gurpreet Singh[1], Pratik Kumar[2], Sanjay Thulkar[3]

1 & 2, Medical Physics Unit, IRCH, All India Institute of Medical Sciences, New Delhi, INDIA
3 Department of radiodiagnosis, IRCH, All India Institute of Medical Sciences, New Delhi, INDIA

ABSTRACT:
Computer aided visualization is a new and evolving topic in the realm of radiology. Digital image processing in case of breast cancer radiology is an area characterized by the need for extensive experimental work to establish the viability of proposed solution to a given problem. The objective of this study is to examine the use of computed fuzzy segmentation technique to enhance the edges of breast lesions in radiological images. This would help the radiologists in their interpretation of medical images. Our major concern is to identify tumor in an image and we have seen that the tumor is clearly recognizable by its texture information.

Key Words: Quality, breast cancer, mammography, screening, radiological images

INTRODUCTION

Breast cancer is reported to be commonest cancer with an annual age adjusted incidence rate of 22-28 per 100,000 women per year in urban areas and 6 per 100,000 women per year in rural areas. Over 75,000 new cases of breast cancer are reported to occur in India per year. Majority of breast cancers in India (50-70%) present with locally advanced disease [1]. Currently available non invasive diagnostic techniques help the radiologists to visualize the suspected lesions with more clarity. However when simple objects are detected on uniform background, the human observer is limited only by noise in the image and this problem is scarier in low contrast situations. Computer analysis of images can begin only in digital format. Most of the image analysis algorithms perform segmentation as a first step towards producing description. This technique basically divides the spatial domain. There exists no explicit segmentation algorithm which will work on all images. So algorithm has to be developed especially for the problem in hand. Segmentation involves the edge detection approach. Detecting edges is a basic operation in image processing. The edges in an image hold much of operation in the image [2-3].

Texture analysis has been used by some researchers to separate out the differently textured regions [4-5]. In one study, a method was developed and described for texture analysis particularly applicable to radiographic images which seeks to classify each pixel of an image [6]. The Laws method uses filter masks to extract secondary features from natural micro-structure characteristics of the image (level, edge, spot and ripple) which can then be used for segmentation or classification. Most segmentation techniques have been applied to the simple case for which the definition of "homogeneous" is based on a single gray-level characteristic. These methods are usually classified as Thresholding, edge detection or gradient-based [7-8]. Many people have used region growing for segmentation of suspicious portions from the mammograms [9]. In their studies, if pixel intensity falls in a specific range, the region-growing algorithm is applied and the intensity gradient is computed to test whether the candidate pixel satisfies the mean and variance criteria. The problem of this algorithm is that it requires many user-input variables. These variables are actually image dependent and should be determined automatically. Some researchers have used hierarchical (or pyramidal) schemes [10-11]. Some of these techniques have been extended for the use of multiple features and have been applied to the more difficult problem of segmentation based on texture. A natural generalization of applying thresholds to a single gray level characteristic is segmentation based on clustering. Several edge oriented methods have been proposed, these generally attempt to locate texture edges based on the computation of a multifeature gradient-like operator [12-13]. Hierarchical approaches using pyramid node linking or applying the split-and-merge algorithm to the co occurrence matrix have also been described [14-15]. In this paper we adopted a texture segmentation technique to find the boundary of the tumor. This method proved to be very successful in detecting the masses in mammograms. In our experiment, we have taken images of mammogram

from Mammographic Image Analysis Society (MIAS) database [16].

This paper enumerates our experience in medical image analysis in general and mammography particular which poses a far tougher challenge. The computer speculated lesion detection algorithm operates on digitized images. The present work involves the mammography of the breast for digital image processing. This work describes the simple approach towards the segmenting the suspicious masses in breast cancer images.

MATERIAL & METHODS

The computer algorithm operates on digitized Mammography images. Thresholding is the process of classifying each pixel in the image as either object pixel or the background pixel. Every pixel value is compared with the threshold value which is dependent on many factors. If the pixel value is greater than the threshold then the pixel is labeled as object pixel otherwise it is labeled as background pixel.

Thresholding [3] may be viewed as an operation that involves tests against a function T of the form

$$T = T[x, y, p(x, y), f(x, y)] \qquad (1)$$

where f(x,y) is the gray level of the point (x,y) and p(x,y) denotes some local property of this point-for example, the average gray level of a neighborhood centered on (x,y).

A threshold image g(x,y) is defined as

$$g(x, y) = \begin{cases} 1 & if \quad f(x, y) > T \\ 0 & if \quad f(x, y) \le T \end{cases} \qquad (2)$$

When T depends only on f(x,y) (that is, only on the gray level values) the threshold is called global.If T depends on both f(x,y) and p(x,y), the threshold is called local. If ,in addition,T depends on the spatial coordinates x,y the threshold is called dynamic or adaptive. Segmentation is accomplished by scanning the image pixel by pixel and labeling each pixel as object or background, depending on whether the gray level of that pixel is greater or less than the value of T. Here the value of T is calculated by different ways. For instance, the value of T which partitions the histogram well is taken as threshold.

To convert the spatial domain image into the fuzzy domain, we consider the spatial arrangement of gray levels of pixels over a window. The fuzzy property can be expressed in terms of a membership function. A membership function to this effect is defined by the Sigmoid function.

$$\mu_{(k,l)}(i, j) = \frac{1}{1 + \exp\left[-\{(x(i, j) - x(k, l))/\tau\}^2\right]} \qquad (3)$$

Where x (i, j) is the gray level of the pixel at the (i,j)th position and τ is the fuzzifier which is taken as the window size. In our experiment, we take the value of τ as 5. We note that

$$\mu_{(i,j)}(k, l) = 0.5 \text{ if } x(i, j) = x(k, l) \qquad (4)$$

To consider the response from the neighboring pixels, we obtain the cumulative response of (i, j)th pixel as follows

$$I(i,, j) = \frac{\sum_{k,l} \mu_{(k,l)}(i, j) * x(k, l)}{\sum_{k,l} \mu_{(k,l)}(i, j)} \qquad (5)$$

This is the defuzzified response of the (i,j)th pixel over the window of size 5. This process is repeated for all pixels in the image and this yields a texture image consisting of all defuzzified values. For convenience of notation, we designate the matrix formed by I(i, j) as the response matrix.

The input image is preprocessed using the equation

$$\mathrm{I(x, y)} = \frac{/}{w^2}\{\sum_{k=i-M}^{i+M} \sum_{j-M}^{j+M} I^2(k,l) - \frac{1}{w^2}[\sum_{i-M}^{i+M} \sum_{j-M}^{j+M} I(k,l)]\} \qquad (6)$$

Where the w is window size. The linear operation of interest consists of multiplying each pixel in the neighborhood by a corresponding coefficient and summing the results to obtain the response at each point. Here, the neighborhood is of size 5 x 5 which gives the best result, so 25 coefficients are required. These coefficients are arranged as a matrix called a window. I(i, j) is the input image and M= (w-1)/2. Here (i, j) is the position of pixel in the image. ‘g’ is the various gray levels which vary from 1 to 256 of digital image after preprocessed from Eq. 6. Gray levels are derived using the histogram ‘hist(g)’[17]. The priori probability is computed using the equation

$$Pi(T) = \sum_{g=a}^{b} hist(g) \qquad (7)$$

ISSN: 0974 – 3987
IJBST (2008), 1(3):10-16

Where the mean and standard deviation are computed using following equations

$$\mu_i(T) = [\sum_{g=a}^{b} hist(g)g] / Pi(T) \qquad (8)$$

$$\sigma_i^2(T) = \left[\sum_{g=a}^{b} h(g)g^2\right] / \left[Pi(T) - \mu_i^2(T)\right] \qquad (9)$$

Then minimize the criteria function to obtain the threshold

$$R(T) = 1 + 2[P_0(T)\ln\sigma_0(T) + P_1(T)\ln\sigma_1(T)] \quad (10)$$

Where the Po (T) corresponding to the probability of threshold 1 to 100. For Po (T), we have used a=1 and b=100. P_1(T) is the probability corresponding to a=101 to b=255. The thresholds values vary from image to image in mammography. For mammography it is around 125 - 185 which satisfies most of breast cancer images.

The images were preprocessed using MATLAB function (MATLAB version 7.01). The object to be segmented differs greatly in contrast from background images. Changes in contrast can be detected by operators that calculate the gradient of an image using the Sobel Operator which creates a binary mask using a user specified threshold value. First image is to be processed with equations (5) and (6). For calculating the threshold using the equation (11) the priori probability, mean and standard deviation are derived using the equations (7), (8) and (9). After that depending upon the shape of the segmented image one can use different structuring elements namely line, diamond and disk shaped. In the present work with mammogram of breast we first used disk shaped structuring element having radius four and six units. Also the connectivity was set to four dimensional to remove the diagonal connection. Compared to the original image there are the gaps in the lines surrounding the object in gradient mask. The linear gaps disappear with the sobel operator. The sobel edge detector uses the masks to approximate digitally the first derivatives. In order to make the segmented object look more natural, we made the objects smooth by eroding the image with a diamond structuring element. Diamond structuring element extends $\pm$ 5 pixels along the horizontal and vertical axes.

RESULTS

Figure 1 and 4 are Mammography pre-processing images. Fig 2 is the post processing image of Fig 1 and Fig 5 is the post processing processed images of Figure 4. Figure 4 is taken from the R Gupta [6]. The differences between pre and post processed images are marked and clearly visible. Radiologists have also shown confidence in images after analysis. Histograms of digital images 1 and 4 with total 256 total possible intensity levels in the range [0,255] are given in the Fig. 3 and 6 respectively.

Table 1: Table (N=15) shows the Mean, Std. dev., S.E., CI levels of diameters read by between different groups

Sr. No.	Percent diff. bet. groups	Mean	S.D.	S.E.	95% CI of difference	
					LB	UB
1	A vs. CAD	3.23	2.99	2.19	-1.67	1.19
2	B vs. CAD	0.09	4.13	0.82	-1.79	1.61
3	C vs. CAD	2.22	4.97	1.79	-5.93	1.47
4	A vs. C	0.19	6.13	1.82	-2.49	4.61
5	B vs. C	3.14	5.77	2.07	-1.47	5.93
6	A vs. B	0.89	3.03	0.80	-0.76	2.56

Fig. 1 Mammography pre-processing image

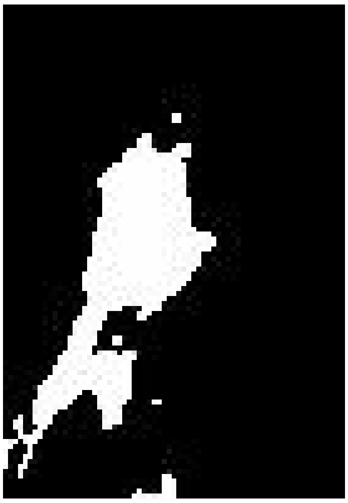

Fig. 2 Mammography post-processing image

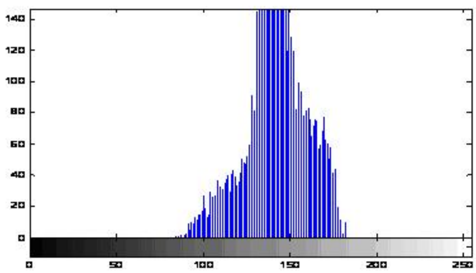

Fig. 3 Histogram of digital image 1 with total 256 total possible intensity levels in the range [0, 255].

Fig. 4 Mammography pre-processing image

Fig. 5 Mammography post-processing image of Fig. 4

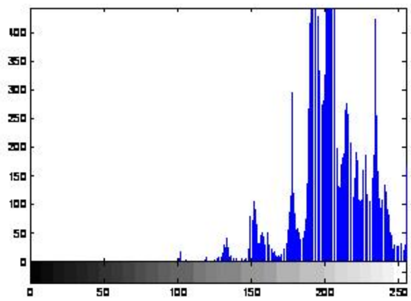

Fig. 6 Histogram of digital image 4 with total 256 total possible intensity levels in the range [0, 255].

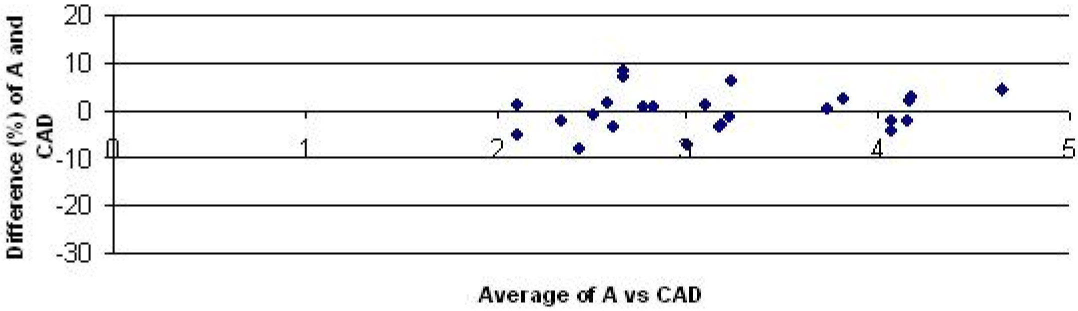

Fig. 7: Graph representing percent difference and averages between A vs. CAD.

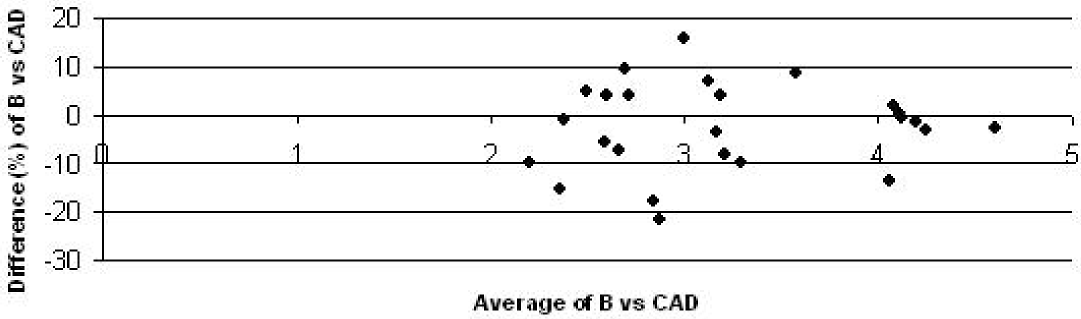

Fig. 8: Graph representing percent difference and averages between B vs. CAD.

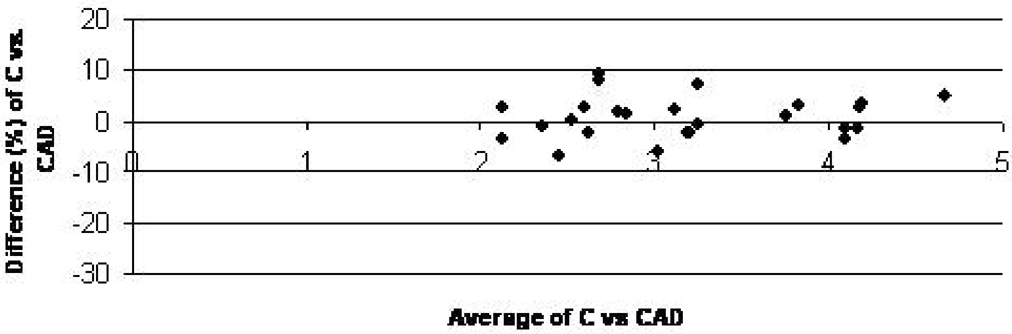

Fig. 9: Graph representing percent difference and averages between C vs. CAD.

The percent difference between the two observers measurements (maximum diameters on the images) are plotted against the average of the both measurements which is the Bland & Altman plot. Diameters are measured in the centimeter (cm). This Bland & Altman plot is a statistical method to compare maximum diameters on same medical images as marked by different radiologists to see the confidence interval among the measurements at 95% level. In our work there are three observers which measure the dimensions on raw images i.e. without processing. Also they measured dimension after processing images. Where A corresponds to group of medical images diameters measured by radiologist I, B corresponds to group of medical images whose diameters are measured by another radiologist, C corresponds to the measurement of 3rd radiologist and CAD is refer to group

whose diameters are measured from computer processed images accepted by all radiologists. All radiologists are having the experience of several years. In the present work there are total 15 images used for study. In Fig. 7 and Fig. 8 graphs between percent difference and averages of A vs. CAD and B vs. CAD have been drawn. In Fig. 9 graph between percent difference and averages of C vs. CAD have been drawn. Table 1 shows the Mean, standard deviation (S.D.), standard Error (S.E.), Confidence interval (CI) of difference at 95% levels between different groups.

DISCUSSION AND CONCLUSION

In the present work, a computer-aided system for segmentation of masses is presented. In our experiment, we have taken images of mammograms which have a dominant texture for classification. We attempted to segment the image based on texture only. Our major concern is to identify tumor in an image and we have seen that the tumor is clearly recognizable by its texture information. The confidence interval in the means of difference is minimum between the groups read by A vs. CAD and B vs. CAD (where CAD is computer based post processing dimensions) as given in Table 1 which means the best accuracy for screening the mammograms is also given by dimensions of post processing mammograms in these groups. Minimum the difference between the values in the confidence interval at 95% the better the consistency for those data values. But on the other hand we observed in the results of other groups that confidence interval in the means of difference at 95% level is maximum between the groups read by A vs. C, B vs. C and A vs. B groups. The diameter measurements made by radiologist B is far less than the radiologists A and C. The differences (%) among A vs. CAD lies mostly within mean ± 1.96 SD, the two methods may be used interchangeably. So, CAD can be used for dimensional measurement. Similarly, the differences (%) among B vs. CAD & C vs. CAD also lie mostly within mean ± 1.96 SD; any method may be used interchangeably. So, CAD performs well in the dimensional measurement.

The study demonstrated that computer analysis can help the radiologists substantially in their screening efficiency. In this paper we used a statistical algorithm which partitions the digital mammography images in to homogeneous texture regions. The masses are generally recognizable by texture. In this work, we attempted to segment mammogram containing masses by its texture. The fuzzy technique is explored in this work. A membership function to this effect is defined by the Sigmoid function. The defuzzified image is segmented using thresholding. Our major concern is to identify tumor in an image and we have seen that the tumor is clearly recognizable by its texture information. The detection method proposed here would take in to the consideration the goal of the proposed computer aided diagnosis system. For a detection system, the important issue is to alert radiologists to suspicious areas on the mammogram. Computer vision techniques have the distinct advantages of being as reproducible as the underlying computer code on which they are based. In the present work we have also used the sober filter based on mathematical aspects which smoothes the image to different extents in the direction of intensity gradient [18]. Computerized detection ability is impervious to the human observer's (radiologists) fallibility which may occur in a busy clinical department.

Further work may be done to improve the accuracy and reproducibility of radiological images further so that it may help the human interpretation of images. The paper indicates the untapped potential of image analysis techniques in routine use of clinical radiological imaging. In this way the entire fields of image processing and computer vision open up to yield interesting techniques for better visualization of medical images [19]. The increase in quality performance is of great value to the radiologist, who often works in a busy environment. Most importantly the further approach and future work which has to be done involves the development of Computer Aided Diagnosis for early detection of breast cancer and even the classification of suspicious and non-suspicious masses.

REFERENCES

[1] Chopra R. (2001) The Indian scene, *J clinical onc,* vol. 15, pp. S106-111.

[2] Hoskins PR., WN. McDicken (197) Colour ultrasound imaging of blood flow and tissue motion, *British Journal of Radiology,* Vol. 70, pp. 872-892.

[3] Gonzalez R. C., R. E. Woods and S. L. Eddins (2005) *Digital Image Processing.* Delhi: Pearson Education publishers.

[4] Hanmandlu M., V. Krishna and S. Vasikarla (2004) A Fuzzy Approach to Texture segmentation, In proceedings of the International Conference on Information Technology IEEE, 2004, pp. 636-642.

[5] Chalana V., Y. Kim (1997) A Methodology for Evaluation of Boundary algorithms on Medical Images, IEEE Transactions on Medical Imaging, vol. 16, pp. 642-652.

[6] Gupta R., PE. Undrill (1997) The use of texture analysis to identify suspicious masses in mammography, *Phys Med Biol,* Vol. 15, pp. 835-855, 1997.

[7] Pal NR., SK. Pal (1989) Entropic Thresholding, *Signal Processing,* Vol. 16, pp. 97-108.

[8] Behrenbruch CP., S. Petroudi (2004) Image filtering techniques for medical image post processing, *British Journal of Radiology,* Vol. 77, pp. 126-130.

[9] Weska JS. (1978) A survey of threshold selection techniques, *Computer Graph and Image Processing,* Vol. 7, pp. 259-265, 1978.

[10] Perkins WA. (1980) Area segmentation of images using edge points, *IEEE Trans Pattern Anal Mach, Vol. 2, pp. 8-15.*

[11] Coleman GB., HC. Andrews (1979) Image segmentation by clustering, in *Proceedings IEEE,* 1979, pp. 773-785.

[12] Grinaker S. (1980) Edge based segmentation and texture separation, in *Proceedings of the 5th Int Conf Pattern Recog,*1980, pp. 776-780.

[13] Bezdek JC., L. Hall (1993) Review of MR image segmentation using pattern reorganization, *Med Phys,* vol. 20, pp. 1033-1048, 1993.

[14] Pietikainen M. , A. Rosenfeld (1981) Image segmentation by texture using pyramid node linking, *lEEE Trans Syst,* vol. 11, pp. 822-825, 1981.

[15] Chen PC., T. Pavlidis (1979) Segmentatlon by texture using a co occurrence matrix, *Comput Graph Image Processing* vol. 10, pp. 172-182, 1979.

[16] Suckling J. (2003) Mini-MIAS (Mammographic Image Analysis Society), http://peipa.essex.ac.uk/info/mias.html

[17] Selvathi D., SS. Thamarai (2006), Tumor detetction in brain magnetic resonance images using modified thresholding techniques, *in Proceedings IEEE,* pp. 327-332.

[18] [18] Abutaleb AS. (1989) Automatic thresholding of gray level pictures using two dimensional entropy, *Comput Vision Graphics Image Processing,* vol. 47, pp. 22-32, 1989.

[19] [19] Huber S., M. Michael (2000) Computer assisted semiquantitaive analysis of color Doppler analysis, *J Ultrasound Medicine,* vol. 19, pp. 601-606.